DOCUMENT RESUME

ED 407 984 JC 970 324

AUTHOR Cowin, Bob
TITLE Fourth Generation Evaluation, Program Review and the Institutional Researcher.
PUB DATE May 96
NOTE 27p.; Paper presented at the Annual Forum of the Association for Institutional Research (36th, Albuquerque, NM, May 5-8, 1996).
PUB TYPE Information Analyses (070) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Case Studies; College Planning; Community Colleges; *Evaluation Criteria; *Evaluation Methods; Foreign Countries; *Institutional Research; Organizational Objectives; *Program Evaluation; Research Utilization; *Self Evaluation (Groups); Two Year Colleges
IDENTIFIERS AIR Forum; *Fourth Generation Evaluation Theory; *Researcher Role

ABSTRACT

Program evaluation can be understood as the process of looking at how all aspects of a program or department have been functioning as the basis for informed planning and decision making. Although the objective dimensions used in evaluations can vary, methodologies can be categorized according to the four category framework (i.e., describing reviews in terms of goal-based, responsive, decision-making, and connoisseurship models) or the external-internal-self study continuum, which categorizes models according to the "distance" of the evaluators from the institution. Emerging approaches to evaluation tend to view evaluation as a process rather than a discrete event and include internal evaluation methods and the controversial fourth generation evaluation. This approach suggests that the first three generations of review (i.e., measurement, description, and judgment) fail to accommodate value pluralism, extend too much power to managers, and are overly dependent on the scientific paradigm of inquiry. Fourth generation review applies a constructivist world view, denying the existence of a scientifically verifiable reality, to a responsive model of evaluation, in which all stakeholders' concerns are taken into consideration. Although the role of the institutional researcher varies with the evaluation methods used, the emerging models call upon researchers to provide knowledge of institutional contexts and decision-making in addition to technical data. Includes a description of a modified fourth generation review undertaken at Douglas College. Seventeen principles related to unit review are appended. (HAA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 407 984

Fourth Generation Evaluation, Program Review and the Institutional Researcher

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Cowin

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented at the Annual Forum of the
Association for Institutional Research
May 1996

Bob Cowin
Institutional Research Office, Douglas College
New Westminster, BC
Canada V3L 5B2

JC 970 324

BEST COPY AVAILABLE

Abstract

This paper highlights the wide range of approaches to conducting program reviews, provides some frameworks for categorizing them, and then focuses on internal evaluation and the emerging and controversial "fourth generation" evaluation advocated by Guba and Lincoln. To the extent these new approaches are adopted, institutional researchers must move beyond technical knowledge to what Terenzini calls "issues and context intelligence". A description of a modified fourth generation review system illustrates how the responsive constructivist paradigm is influencing evaluation practice in a community college.

Contents

Introduction 1

Program Evaluation 2

- Definitions 2
- Diagnostic Questions to Describe Review Systems 3
- Frameworks for Categorizing Program Review Methodologies 4
  - Four Category Framework 4
  - External -- Internal -- Self Study Continuum 5
- New Approaches 6
  - Internal Evaluation 7
  - Fourth Generation Evaluation 8
  - Relationship to the Learning Organization 11

Fourth Generation Evaluation and the Institutional Researcher 12

- Nature of Institutional Research 12
- Institutional Research's Contribution to Program Evaluation 13

Example of a Modified Fourth Generation Review System 15

- Description of Unit Reviews 16
- Commentary 18

Conclusion 19

Appendix: Unit Review Principles

Introduction

This paper summarizes several general approaches for reviewing programs, giving particular attention to recent developments in the field of program evaluation. Despite the emphasis on emerging schools of thought, and especially on "fourth generation evaluation," it does not argue that any single approach is best. Rather, different methods have different strengths and weaknesses, and are thus appropriate in different situations or for different purposes. What is important is for the practitioner to be aware of a variety of approaches, and their implications, in order to choose among them explicitly -- something which institutional researchers are not always well equipped to do.

After touching on a few definitional matters, the first section of the paper presents two frameworks for categorizing evaluation methodologies. Within each framework, an approach which has grown in popularity during the 1980s is highlighted. These emerging approaches, not surprisingly, parallel the thinking in the "new" style of management: collaborative, empowering, systems oriented and respectful of diverse points of view.

The second section briefly discusses the role of the institutional researcher in program review: the contribution he or she can make to the process as well as the benefits to institutional research from participating in the evaluation enterprise. It is written from the perspective of what Terenzini (1993) has described as the "organizational intelligence" conception of institutional research.

The concluding section illustrates what the theory can mean in actual practice. It describes how some of the recent thinking in the program evaluation literature is being applied in a community college. Referring neither to an exemplary organization, nor to an application which is appropriate in all situations, the example nevertheless illustrates one of the newer approaches to program review. It demonstrates the tradeoffs one makes in choosing among methodologies.

The paper is a work of synthesis. Little, if anything, is new. Topics are treated briefly and superficially. It is intended as an overview of a field in which road maps are lacking, practitioners are often unaware of the paradigms within which they operate, and theorists tend to write for an academic audience which shows little appreciation of the needs of the average, typically untrained, practitioner.

Program Evaluation

Definitions

Program "review" is used here synonymously with program "evaluation", and "program" is used in the generic sense to include support services as well as instructional programs. (In public administration, a program is a general effort that brings together people and activities toward some, perhaps poorly, defined and financially supported goals. Thus postsecondary institutions have three main types of programs: instructional programs, programs which provide support services directly to students, and programs which indirectly support the institution's mission by providing services for employees, e.g. the financial services of the Purchasing Office.)

A number of definitions of program evaluation exist, but I prefer an informal description which faculty and other employees can readily grasp and support: program evaluation is the process of looking at how all aspects of a program or department have been going as the basis for informed planning and decision making. The key feature is that program evaluation is comprehensive, distinguishing this type of evaluation from specialized reviews such as curriculum review, assessments of organizational climate, or financial audits. (Each of these more specialized reviews could be a component of a program evaluation.) Although comprehensive, program reviews are not all-encompassing; operational details, for example, are beyond the scope of program evaluation.

Program evaluation concerns the role and performance of groups of people. While the intended role of individuals may be examined, the actual performance of individuals is not to be considered. The performance of individual employees falls within the domain of personnel evaluation, while the performance of individual students is assessed by faculty as part of their instructional duties.

Program reviews can be formative or summative. The distinction here is purpose, not technique. Scriven (1991) reports Robert Stake's illustration of the difference: "When the cook tastes the soup, that's formative; when the guests taste the soup, that's summative." Except in times of budget crisis or enrolment problems, postsecondary institutions rarely engage in truly summative evaluation. Most program evaluation in the educational environment is formative.

Diagnostic Questions to Describe Review Systems

In contrast to a remarkably similar repertoire of techniques for delivering instruction and facilitating learning, the variety in the ways by which institutions evaluate their programs can be overwhelming. The following checklist describes only a number of the objective dimensions by which methodologies vary, setting aside more the elusive dimensions such as purpose or epistemology:

Units Evaluated:
Instructional areas only or service departments as well?

Unit of Analysis:
Individual disciplines and units, or several related entities combined?

Frequency:
Regularly scheduled or as needed? Annually or every few years?

Intensity:
Overview or in-depth study?

Scope:
A full range of topics or focused?

Methodological Flexibility:
Extent of standardization?

Duration of Review:
Weeks, months or semesters?

Committee membership:
Mix between internal and external membership?

Primary Audience:
Local, institutional, or external?

Support Services:
Amount of centrally provided data-gathering and procedural assistance?

With so much diversity in program review systems across institutions, and sometimes across programs within the same institution, it helps to be able to group methodologies into a few basic categories. The following section presents two simple frameworks for categorizing evaluation methodologies and purposes.

Frameworks for Categorizing Program Review Methodologies

Four Category Framework

The Association for the Study of Higher Education sponsored a review in the mid eighties of the literature on program review in higher education (Conrad and Wilson, 1985). Although it is now becoming a little dated, the report provides a good introduction to program evaluation "approaches, expectations and controversies." It concluded that program reviews in colleges and universities generally fall into one of four approaches:

1. Goal-based model

> The oldest and what in 1985 was the most widely used model of academic program review, the goal-based model assesses how the program has performed in relation to what was intended. It is typically used to make predominantly summative decisions about questions such as resource allocation and program continuance. The methodologies tend to be quantitative.
>
> The chief limitations of the model are the specification of goals (not all goals are written down and the temptation to identify a goal for every conceivable desire can become obsessive), insensitivity to outcomes that are unrelated to goal statements, the assumption that valid measures can be found for all goals, and the inflexible way in which *a priori* goals drive the evaluation process.

2. Responsive model

> Finding its roots in the goal-free model of Scriven (1972, 1973), the model argues that unintended or side effects are often as important as intended effects. It is organized around the concerns and issues of stakeholding audiences, i.e. it investigates what various constituents believe a program is accomplishing and their concerns about the program. The approach tends to be formative and to use qualitative methods.
>
> The responsive model is at the heart of the fourth generation evaluation discussed below.

3. Decision-making model

> This model explicitly attempts to link evaluations with decisions that have to be made. Structured to meet administrative information needs, it strength lies in ensuring the results are utilized. It tends to be summative and to use quantitative methodologies.
>
> The principal criticism is that this model accepts the decision context and values/criteria that have been defined, perhaps incompletely or inaccurately, by the

decision makers. It assumes rational decision making and that all decision alternatives can be accurately anticipated. Sometimes it is difficult to specify exactly who are the decision makers in complex organizations where decisions are made jointly at several organizational levels.

4. Connoisseurship model

Evaluations conducted using this approach rely heavily on the perspectives and judgments of experts, i.e. of individuals with superior knowledge and a commonly shared value system. Emphasis is placed on the person chosen to conduct the evaluation as no two experts will have the same value structures or weigh criterion equally. Lacking methodological guidelines, although typically qualitative and formative, and placing a premium on the evaluator's judgment, it is hard to verify whether the evaluator's perceptions are accurate.

Since Conrad and Wilson were published in the mid 1980s, a form of responsive evaluation has attracted considerable attention and debate: fourth generation evaluation (see below).

External -- Internal--Self Study Continuum

Another simple way of categorizing evaluation models is according to the "distance" of the evaluator(s) from the activities being assessed:

1. External evaluators
   - external to the organization, as well as to the program

2. Internal evaluators
   - employees of the organization who are not directly involved with the program, but who are familiar with the culture and environment of the program

3. Self study
   - program employees step back from day to day activities to assess their program

As with the four category framework, each approach within this framework has its distinctive set of strengths and weaknesses:

1. External evaluation

Evaluation by third parties is especially appropriate for summative purposes or where there are financial concerns such as fraud or waste. External personnel can bring specialized knowledge, have sufficient distance to see the big picture, and provide labour for the extra work associated with an evaluation. Whether external evaluators are

BEST COPY AVAILABLE

actually independent and objective is sometimes a moot point, but they can bring a fresh perspective and provide credibility for the evaluation.

Traditional external evaluators review the extent to which the program is achieving its goals for the purposes of accountability, resource allocation, control and program justification. As a consequence, external evaluation can engender defensiveness and anxiety within the program. This, in turn, can hinder the implementation of recommendations where the success of implementation depends in part on the goodwill of program employees. To the extent that a goal-based model of evaluation is used, the limitations of that approach are also present in external evaluation.

2. Self study

The objectivity of self assessments, the polar opposite of external evaluation, is always suspect. Even with the best of intentions, it can be difficult for program employees to break with historical practice and viewpoints. With less than goodwill, cover-ups and biased data collection are all too easy with self assessments.

The preeminent advantage of self study approaches is the ownership or support for its recommendations by those delivering the program. If major changes are needed, and the program managers and employees reached that conclusion themselves, then they become advocates for change rather than potential resistors. The evaluation process itself becomes part of the implementation strategy.

Self studies are used for formative purposes. When done well, they are proactive and educational. When done poorly, they are self-serving and not credible.

3. Internal Evaluation

Internal evaluation, occupying an intermediate position on the external audit/ self study continuum, combines a modest amount of "distance" (helpful for identifying problem areas and putting them in context) with modest "savvy" as to what form recommendations should take to enhance the probability of successful implementation over the long term.

Internal evaluation is discussed in more detail below. Just as fourth generation evaluation has emerged in the past decade, internal evaluation has come into its own relatively recently.

New Approaches

The program evaluation literature is moving away from suggesting that evaluators must be, or even can be, detached observers or discovers of "truth". Approaches that see evaluation as a process are

emerging to complement methods which are discrete projects or events. Two such approaches are described here.

Internal Evaluation

What was new or distinctive about internal evaluation was the notion that evaluation does not have to be done exclusively by external people. It has really only been since the late 1980s that there has been a huge shift to internal evaluation, with the corresponding acceptance of the approach's legitimacy.

This section draws extensively on Love (1991) and a workshop he gave in 1993 for the American Evaluation Association at its annual conference -- the conference's theme being "Empowerment Evaluation." Love sees internal evaluation as:

> ....a form of *action research*...that supports *organizational development* and *planned change*.... Consequently those responsible for internal evaluation are responsible not only for analyzing problems and offering recommendations, but also for correcting difficulties and implementing solutions.... The study method is tailored to the constraints posed by political and practical considerations....

The desire in internal evaluation is to foster a learning culture within the organization to make the program more effective, rather than just to judge whether it worked. Thus the evaluator(s) in an internal system is more involved with the program under review than in the self-contained, project oriented model of evaluation. The theory is that because of the long term commitment and continuity of the individuals involved in the evaluation, internal evaluation can encourage better communication across the organization and greater utilization of evaluation information.

In internal evaluation, the perspective changes from "Are goals being met?" to "Are we the best we can be?" Carrots (incentives) are used rather than sticks (punishments) as the internal evaluator seeks to serve as a change agent. This shift in the tenor of program evaluation parallels trends in management practices in general:

Old Style Management	New Management
Close control	"Hands off"
Supervisor	Catalyst
Expert	Resource person
Director	Coach, problem solver
Controller	Networker, champion

Whereas the external evaluator or auditor can be threatening to program personnel, eliciting defensiveness and fear about hidden agendas and the assignment of blame, the internal evaluator seeks to affirm and renew by giving voice to stakeholders, making work visible, clarifying, providing new information and making recommendations for improved performance. It is a very different face to evaluation than the reactive, diagnostic, isolated face of some traditional approaches.

Fourth Generation Evaluation

Whether or not one agrees with the approach advocated in it, Lincoln and Guba's 1989 book, *Fourth Generation Evaluation*, is required reading for anyone interested in program evaluation. The authors see three periods in evaluation history:

First Generation: Measurement
- 1900s to 1930s
- precursor to program evaluation in that it looked only at individuals and not at groups of people
- development of mental tests to screen military personnel and to measure various attributes of school children

Second Generation: Description
- 1930s to 1950s
- described patterns of strengths and weaknesses with respect to stated objectives
- evaluating non-human elements (e.g. programs, strategies, organizations) as well as human ones, it represents the birth of program evaluation

Third Generation: Judgment
- 1960s and 1970s
- required that program objectives themselves be taken as problematic. Goals, no less performance, were subject to evaluation
- the issue of values that are inherent in all evaluations became explicit in that clear standards were needed against which judgments are made

Guba and Lincoln's criticisms of the first three generations of evaluation are also threefold:

1. Failure to accommodate value-pluralism
   The question of whose values will dominate in an evaluation is often overlooked.

2. Tendency to managerialism

To the extent managers have ultimate power to determine what questions the evaluation will pursue, how the answers will be collected and interpreted, and to whom the findings will be disseminated, the conclusions are in part predetermined and other stakeholders are disempowered.

3. Over commitment to the scientific paradigm of inquiry

The over dependence on formal quantitative measurement implies that what cannot be measured does not matter. Furthermore, it leads to "context stripping", resulting in the truncating of information.

Lincoln and Guba's response to these criticisms is to recommend a fourth generation of evaluation: responsive constructivist evaluation.

Fourth generation evaluation emerged from the application of a constructivist world view to the responsive model of evaluation mentioned above. Lincoln and Guba describe responsive evaluation as follows:

> "Responsive evaluation was so named by its originator, Robert Stake, to signal the idea that all stakeholders put at risk by an evaluation have the right to place their claims, concerns and issues on the table for consideration (response), irrespective of the value system to which they adhere. It was created as the antithesis of preordinate evaluation, which assumes the evaluator and the client together possess sufficient information and legitimacy to design and implement an evaluation, without the need to consult other parties...."

The jump from the position that all stakeholders should have their points of view acknowledged in an evaluation to constructivism is a small but significant one.

Constructivism denies that there is an objective, scientifically verifiable reality for humans to discover. Rather, there are only "constructs" which exist in the mind: as people experience sensations and gain data, they interpret these events in the context of their previous knowledge. The new information, i.e. their interpretations of data, is then added to their previously constructed body of knowledge. Thus "knowledge" is not the objective reality of positivism but rather the meanings and interpretations of individuals striving to make sense out of their situations.

The nature of a construction depends on two things:

- the range of information available

- the constructor's sophistication in dealing with that information.

Constructions are formulated and reformulated on the basis of available information and the individual's ability to rearrange information items in new patterns.

The methodological implications of constructivism are compatible with a political environment but not with a technological, hierarchical environment. Because constructivists do not believe unbiased, objective answers exist to any question in a program evaluation, they seek instead to build consensus among parties as to what needs to be done; negotiation is a key dynamic. When constructivism is merged with the value pluralism of responsive evaluation to form fourth generation evaluation, the goals of evaluation include bringing about change and empowering all stakeholders.

Constructivists deny an evaluator can stand outside what is being evaluated as a neutral observer, and therefore reject controlling, manipulating (experimental) methodologies in favor of naturalistic inquiry. They explicitly acknowledge that the findings of a study exist because of an interaction between the observer and the observed, with different interactions or different observers yielding different findings. This orientation towards qualitative research has implications for institutional researchers and the way we go about gathering information for decision makers:

> "Generally, quantitative methods are focused on obtaining specific items of information, and great stress is placed on systematic approaches across people or places; qualitative methods are more likely to involve techniques that broaden the information base -- adding more perspectives, raising additional issues, constantly accumulating more details, with less attention to inconsistencies in data collection procedures." (Hedrick, 1994)

The fourth generation evaluations advocated by Lincoln and Guba are conducted in such a way that each stakeholder group must confront and deal with the constructions of all others. As each group copes with the perspectives of others, their own views alter by virtue of being better informed. Some consensus will be achieved, but conflicts will remain whose resolution requires the introduction of outside information (which it becomes the evaluator's task to obtain). With as much of this information as is feasible to collect, the evaluator prepares an agenda for negotiation. Further consensus is reached. In remaining areas of conflict, stakeholders at least understand what the conflict is about and where other groups stand in relation to it.

This approach can be summarized in four steps:

1. identify stakeholders
2. claims, concerns and issues raised by each stakeholder group are introduced to all other groups
3. those claims, concerns and issues that have not been resolved become the advance organizers for information collection by the evaluator (the information may be quantitative or qualitative)
4. negotiation among groups in an effort to reach consensus on each disputed item.

It involves giving control of the evaluation process to stakeholders and a high tolerance for ambiguity.

Relationship to the Learning Organization

Why would an institution consider adopting one of the newer approaches to program review, be it internal evaluation or fourth generation evaluation? One reason might simply be that every approach has its limitations and that new approaches are tried as part of the long-term cycle of seeking reviews with less critical limitations. Another reason is that to the extent the organization implicitly or explicitly accepts the notion of the learning organization, the newer program review methodologies can help transform institutions into such organizations.

Learning organizations, as the term is used by Senge (1990) and others, are ones which have the capacity to adapt and innovate, based on learning from past experience and in anticipation of the future. They support continuous improvement at three levels: individuals, work groups and the organization as a whole.

The learning organization is a response to a rapidly changing environment in which there is no steady state. Every time a strategy is developed or a task performed, it is reflected upon and used as an opportunity for considering how it might be improved. This approach to work is premised on team work, responsibility-taking, self-management and learning. A significant component of this transition to a new style or workplace involves moving from a traditional hierarchy to a self-regulating, team-based organization.

The new program evaluation methodologies reinforce elements, or what Senge calls "disciplines", of the learning organization. Team learning, questioning the taken-for-granted mental models which limit thinking, the emergence of a shared vision from people who talk with each other and

respect each other's viewpoint, planning as a process rather than a single event, looking for underlying patterns rather than specific people or events as the roadblocks to change....all of these are present to varying degrees in contemporary notions of management and evaluation.

Fourth Generation Evaluation and the Institutional Researcher

The role of the institutional research office in program evaluation depends in part on how broadly institutional research is conceived, as well as on the model(s) of evaluation which the institution has adopted.

Nature of Institutional Research

Patrick Terenzini (1993) recounts the uncomfortable silence of a group of institutional researchers in an elevator searching for a quick answer to a hotel guest's question, "What's institutional research?" The answer he proposes is sufficiently broad to encompass the wide range of activities in which (at least some) institutional researchers engage. If this conception of institutional research is sound, then institutional researchers have a role in program evaluation beyond that of providing data upon request.

Terenzini advocates viewing institutional research as organizational intelligence. This term refers to "the data gathered about an institution, to their analysis and transformation into information, and to the insight and informed sense of the organization that a competent institutional researcher brings to the interpretation of that information." (p. 3) He sees three different, but equally important and interdependent, dimensions to organizational intelligence:

1. Technical/Analytical Intelligence

   Technical intelligence includes factual knowledge as well as quantitative and qualitative research skills. It is fundamental, but without the other forms of intelligence it has little utility or value: "By itself, it consists of data without information, processes without purposes, analyses without problems, and answers without questions."

2. Issues Intelligence

   Issues intelligence comprises knowledge about the major decision areas that face the institution. It involves an understanding of how institutions function and how decisions are made, i.e. an appreciation of the political character of decision making. "It is knowledge of how to work successfully with other people (both individually and in groups) to accomplish some goal."

3. Contextual Intelligence

   Contextual intelligence involves understanding the culture both of higher education in general and of the institution. It includes familiarity with the perspectives of students, of college employees, of government and of employers. As with issues intelligence, contextual intelligence provides a basis for judging what is important.

The need for issues and contextual intelligence helps to explain why some technically competent institutional researchers are viewed on their campuses as mildly unintelligible, irrelevant or even counter-productive. It also broadens the role of the institutional research from simply producing and dispensing data to that of brokering information: selectively collecting data from other sources and redirecting it to the right audiences, in the right form, and at the right time.

In order to perform the brokering role effectively, the institutional researcher needs to understand the clients' perspectives and is herself an information customer. The newer conceptions of program evaluation, whether the collaborative, developmental emphasis of internal evaluation, or the negotiating, stakeholder emphasis of fourth generation evaluation, establish an environment in which it is appropriate for the institutional researcher to get to know the people affected by the program under review and how they perceive the program. As dialogue develops, the researcher not only contributes technical know-how towards answering particular evaluation questions, but he or she can share contextual and issues information gathered in other arenas.

Institutional Research's Contribution to Program Evaluation

There is no single paradigm of program review. What works in one place at one time may well not work in a different place or time. Consequently, Barak and Brier (1990) recommend individually developed review processes based on the local environment, history, culture, needs, and the periodic evaluation of the process to ensure that it continues to respond to local needs.

Regardless of the model of evaluation adopted, the technical intelligence of institutional researchers can be valuable for those conducting the review. The institutional researcher already has a wealth of data at his finger tips, already knows the limitations of the data, and has the skills to collect new data efficiently and effectively. Even in the connoisseurship model where visiting experts parachute on to campus on a short fact-finding mission, a visit to the institutional research office can be a valuable part of the itinerary.

In some program evaluation systems, this is about as much of a contribution as the institutional researcher is allowed to make, i.e. the provision of facts and figures upon request. Institutional researchers are exposed, however, to a variety of current and emerging issues in postsecondary education; they understand from environmental scanning the general trends and concerns within the institution and elsewhere in the postsecondary sector. Their work brings them into contact with a variety of stakeholders with differing agendas and points of view.

Institutional researchers are thus well positioned to contribute to the identification of the institutional and broader contexts for program reviews, and to comment on the constraints affecting the implementation of recommendations. Neither internal nor fourth generation approaches to evaluation call upon institutional researchers to pass judgment on the program under review, nor to step out of the support role into a decision making capacity. They do allow, however, for the institutional researcher to be an active resource person to the review, a contributor to all three dimensions of organizational intelligence.

By interacting with stakeholders and evaluators, institutional researchers can learn a great deal about the complexity of the postsecondary world and what conditions are like in the front lines of an institution. The learning from individual reviews may not be very generalizable, but the accumulated information from a number reviews is a rich source of issues and context intelligence -- not the full story, of course, but an important component. It becomes a good way for getting the pulse of the institution.

To the extent that newer models of program evaluation are adopted, the institutional researcher not only has the opportunity to contribute to, and benefit from, the evaluation but may actually be expected

to participate in a significant way. In these models, technical skills may be necessary but they certainly are not sufficient. Institutional savvy, familiarity with the literature of higher education, people skills and a reputation for being trustworthy and a good listener are also necessary.

Example of a Modified Fourth Generation Review System

The paper to this point has been theoretical. This section illustrates what the newer forms of program evaluation can mean in practice by describing an actual application. This application has operated reasonably effectively, but it is far from perfect. Whether it will endure, much less be adopted at other institutions, is open to question. Nevertheless, it does serve to illustrate an approach to program evaluation which differs from traditional approaches.

The case study is of Douglas College, a typical suburban community college of 5,000 full-time equivalent students, the vast majority of whom are located on a single campus. University transfer programs have the largest enrolments, but career/technical programs are also important, especially in health, social services and business. Trades programs are not offered and college preparatory courses tend to be at the upper levels.

Program and service reviews at Douglas College are known as unit reviews. Conceptually, a unit can be any grouping of related activities, but in practice a unit in instructional areas consists of a single discipline or program, e.g. Chemistry or Health Information Services. All support service and non-instructional areas, such as the library and the public relations office, are also reviewed using the same methodology.

Unit reviews were initiated six years ago and over forty have been completed to date. A survey in the autumn of 1995 (60% of respondents were faculty, 40% were administrators, and most had fairly good exposure to at least one review) found reasonable but not universal, support for the system. The principles underlying the unit review system (see appendix) were widely accepted, the process seemed to work not too badly, but follow-up and utilization of the findings at other than the local level were weak -- a

common problem right across the public sector with program evaluations. Some improvements were desired, but there was no consensus as to what is needed. (A strong concern was that a process developed in a benign environment of growth, and predicated on the trust and goodwill of the unit under review, may not be appropriate in the financial downturn the college will experience for at least the next two or three years.)

Description of Unit Reviews

On the external audit/internal evaluation/self study continuum, unit reviews fall somewhere between a self study and an internal evaluation. Consistent with decentralized management and the collegial organization of professionals, two or three individuals from the unit under review compose the core of the review committee. An external perspective is obtained by ensuring at least one committee member comes from outside the unit and, depending on which of the three levels of review is specified, sometimes a couple of additional members from elsewhere in the college or from outside it are present. An institutional researcher serves at the committee's resource person. The unit manager also sits ex officio, keeping a low profile as it is to the manager that the resulting report is submitted.

The review is intended to last about six months and involves anywhere from five to ten meetings of the full committee. Time release is sometimes available for the report writer, depending in part on the number of stakeholder groups to be consulted.

The published generic list of topics to be considered, methodology, and report format for unit reviews are all illustrative, not prescriptive. What is essential, however, is that the perspectives of all stakeholder groups be taken into account and that primary stakeholders be contacted directly through surveys, interviews, meetings or whatever method is deemed appropriate. A stakeholder is defined broadly as anyone who is affected by what the unit does: students, graduates, employers, applicants who were not admitted, employees within the unit, employees who provide or receive service to or from the unit, and so on. Thus the review committee is a fact finding group rather than a panel of experts who already know the views of all stakeholders.

The emphasis on hearing from stakeholders moves the unit review system firmly in the direction of fourth generation evaluation methodology. The principle of obtaining feedback from stakeholders was one aspect of unit reviews which was readily accepted by the College community once it was clear that the purpose of unit reviews is formative and not summative, i.e. geared towards improving whatever currently exists without assessing whether the current situation is adequate or not.

Unit reviews are "responsive evaluation" in that stakeholder input is not restricted to only the topics raised by the review committee or in relation to the stated goals of the unit. Where it differs from the fourth generation methodology of Lincoln and Guba is that stakeholder groups do not provide a second round of feedback in light of what other groups said. Such additional feedback is time and resource intensive, so the unit review committee itself attempts to synthesize and reconcile viewpoints as it reviews the feedback -- not quite the same as fourth generation's call for negotiation, but it does at least allow multiple perspectives to be discussed.

Because each review committee selects its own chair and report writer from among its members, the committee relies on a resource person to help it understand the review process and the flexibility inherent in it. This resource person is an institutional researcher. During the first two meetings, he or she acts as a facilitator as the committee identifies stakeholders and a comprehensive range of potential topics and issues for review. Only after the scope of the review has been broadly determined does the institutional researcher fulfill the more traditional role of information gathering: helping the committee determine which stakeholders could provide input on which topics, and then determining the most appropriate way to gather that input given time and resource limitations, e.g. questionnaire, phone interviews, existing documentation, or focus groups.

Although each review committee custom designs its own data collection plan, the Institutional Research Office holds copies of all survey instruments used in previous reviews. Working mainly with the committee chair, the institutional researcher can bring examples of past questionnaires which addressed similar topics, e.g. organizational climate as perceived by employees or the effectiveness of practicum

arrangements as perceived by supervisors at the work site. This greatly accelerates instrument design while still allowing each review to address the distinctive needs and situation of the unit.

The institutional research office tabulates survey results, using survey software which handles text questions as well as numeric ones. It sometimes conducts interviews or meetings, but generally prefers committee members to perform this role for reasons of time and so that the committee member can receive impressions first hand.

Institutional researchers do not participate in interpreting and drawing conclusions from the data, other than to point out alternative interpretations or to caution where an interpretation is drawn on scant or tenuous data. When recommendations are being made, the institutional researcher may pose questions to help the committee to clarify its intent, or to present information about the institutional context which may help the committee frame its intent with the best possible wording. Even setting aside support units where the report writer may not be a university graduate, a surprising amount of time is spent helping writers structure their report and to appreciate the difference between a review report for decision makers and an academic paper for their colleagues.

Commentary

The unit review system is the product of a number of choices in the late 1980s about the form of program review that best met the needs of the college at that time. A key question was whether to select a model whose strength lies in documenting the conditions in the unit under review. The answer was that this was not the greatest need -- in many instances, a number of problems were already known. What was needed was a way of dealing with the situation, taking into the account the collegial, decentralized structures of postsecondary institutions. The emphasis on unit reviews thus came to be on facilitating change rather than fully describing and judging a unit's condition.

The choice of a flexible methodology to accommodate local distinctiveness makes it more difficult than in standardized approaches to aggregate the findings into a college-wide perspective. Unit

reviews are weak for comparative and summative purposes. The quality of the reviews, and of the resulting reports, vary considerably.

Unit reviews are very dependent on the goodwill and enthusiasm of the committee members from the unit being reviewed. This is a double-edged sword: the reviews are vulnerable to manipulation or whitewash, but have proven to be powerful learning experiences for committee members and an important means of bringing about change where honest appraisals were made in units willing to confront problems. The reviews are slowly modifying an organizational culture which was not especially noted for making decisions based on comprehensive data.

Supporting the unit review process has its frustrations for the institutional research office, not least of which are the large amount of time it requires (in the order of 1 FTE spread across three individuals in the institutional research office) and the normal challenges of committee work. The flexibility of the review system means constantly inventing and doing things for the first time. Workload is difficult to anticipate.

Nevertheless, listening to committee deliberations has been a valuable way to learn about important educational and administrative issues and to gain background context for interpreting data. The institutional researcher sometimes encourages the committee to collect data they had not thought feasible to collect, or relieves them from collecting data because the work, or similar work, has already been done. His or her presence becomes a link into world of information and resources which most committee members are, at best, only vaguely aware. Thus institutional researchers both contribute to, and benefit from, unit reviews.

Conclusion

A wide variety of models of program evaluation are available, each with its distinctive set of strengths and weaknesses, and each useful for particular purposes. The important things are to match the

model with the needs and characteristics of the institution, and to be aware of the paradigm in which the model operates.

Two approaches which have become more widely used in recent years, and more influential in the literature, are internal evaluation and fourth generation evaluation. Neither approach conceives of the evaluator(s) as a detached observer or discoverer of the truth about the area under review. Internal evaluation argues that the evaluator should have some ongoing connections with the reviewed program in order to facilitate the implementation of changes, and to help foster a learning organization. Fourth generation evaluation argues that no single observer can adequately account for all that matters. It claims that the notion of simple objective truth needs to be replaced by an appreciation of the variety in stakeholder constructs and the plurality of values associated with these constructs. Neither internal evaluation nor fourth generation evaluation see program review as a discrete event, but rather as a process in which the information base develops as the review progresses and changes are made.

No matter what model of program evaluation is adopted, institutional researchers can bring issues and context knowledge, as well as technical skills, to the evaluation. The interaction of institutional researchers and evaluators, not to mention the data generated by the review, becomes a source of information about the institution and the postsecondary environment for the institutional researcher.

To the extent the newer forms of program evaluation continue to grow in acceptance and frequency of application, i.e. prove to be the cutting edge or a new paradigm rather than a passing fad, then the role of the institutional researcher in program evaluation expands from a technical orientation to incorporate more of the functions of a coach, consultant, facilitator, and resource person. This is no small challenge, but it is consistent with role of institutional research in other planning and decision making activities.

References

Barak, Robert J. and Breir, Barbara E. (1990) Successful Program Review: A Practical Guide to Evaluating Programs in Academic Settings. San Francisco: Jossey-Bass.

Conrad, Clifton F. and Wilson, Richard F. (1985). Academic Program Review: Institutional Approaches, Expectations, and Controversies, ASHE-ERIC Higher Education Report No. 5, 1985. Washington, D.C.: Association for the Study of Higher Education.

Guba, Egon G. and Lincoln, Yvonna S. (1989). Fourth Generation Evaluation. New Park, California: Sage.

Hedrick, Terry E. (1994). *The Qualitative - Quantitative Debate: Possibilities for Integration*, in Reichardt, Charles S. and Rallis, Sharon F. (1994). The Qualitative - Quantitative Debate: New Perspectives, New Directions for Program Evaluation, no 61 (Spring 1994). American Evaluation Association. San Francisco: Jossey Bass.

Love, Arnold J. (1991). Internal Evaluation: Building Organizations from Within. Newbury Park, California: Sage.

Scriven, Michael.
(1972). "Pros and Cons about Goal-Free Evaluation", Evaluation Comments, no. 3 (December 1972), p. 1 - 4.
(1973) "Goal-Free Evaluation." in School Evaluation: The Politics and the Process, edited by House, Ernest R. Berkley, California: McCutchan.
(1991). Evaluation Thesaurus (4th ed.), Newbury Park: Sage Publications.

Senge, Peter. (1990). The Fifth Discipline: The Art and Practice of the Learning Organization. New York: Doubleday.

Terenzini, Patrick. T. (1993). "On the Nature of Institutional Research and the Knowledge and Skills it Requires," Research in Higher Education, 34 (1), 1-10.

APPENDIX

Unit Review Principles

April 1995

1. Developmental

Unit reviews seek improvement rather than absolute measurement against evaluative criteria. They are not an appropriate process for summative evaluation.

2. Internal Process

The core of the review committee comes from the unit under review. The review is conducted "with" the unit, not done "to" it. It values local expertise and talent.

3. Commitment to Needed Changes

Reviews are designed to help the unit come to its own conclusions about needed changes, rather than to have changes imposed externally. Thus the focus is on a process that helps the unit take ownership of recommended changes.

4. Stakeholder Input

The review actively seeks feedback from all groups who have a stake in what the unit does: unit employees, students, employers of graduates, other units providing service to the unit, etc. The review committee does not necessarily have to act on the feedback it receives, or even agree with it, but it does have a duty to obtain, consider and report it.

5. Collaborative Committee Membership

Some review committee members come from outside the unit and help the committee maintain a perspective which is not unduly influenced by current operational concerns or historical practices. Their role is not to act as advocates for external constituencies, however. All committee members participate equally in decision making, i.e. a collaborative rather than hierarchical approach, although the workload may fall mainly on the report writer.

6. Data Based but Flexible Methodology

The general approach is to collect data from stakeholders and other sources about key aspects of the unit. A variety of quantitative and qualitative methodologies can be adapted to the circumstances and needs of each unit. Data come from sources internal to the unit as well as from external sources. The College assists the committee in gathering data.

7. Action Research

The review is intended to be conducted quickly, collecting sufficient data of sufficient quality to reach defensible recommendations. It examines the results from several data sources to reach conclusions rather than attempting to collect perfect data from a single source. This is a different orientation from academic research.

8. Big Picture

Reviews look at all aspects of the unit, not just areas in which committee members have responsibility or interest. They are concerned with strategic directions, including the goals and mandate of the unit, rather than with operational details and implementation specifics. A holistic view of the unit is sought.

9. Same Principles for All Units

The same basic approach is used, regardless of whether the unit provides instruction or support services. Some topics do not apply in all reviews, e.g. noninstructional areas do not consider curricular matters, but the basic review methodology is similar. A manual describes the methodology.

Please turn over....

10. Not Personnel Evaluation

Unit reviews look at the roles and performances of groups of employees, sometimes the roles of individuals, but never the performance of individuals. The evaluation of individual employees is done through personnel procedures. The evaluation of individual students is done by faculty through instructional procedures.

11. Open Process

Summaries of the review report are distributed to decision making bodies throughout the college, with the full report and appendices available to any employee upon request.

12. Advisory

The review results in recommendations, not decisions. The implementation of these recommendations must be approved through regular decision making processes within the College. Unit reviews are not means of doing "end runs."

13. Reflective Team Learning

Committee members identify issues together, review the same information, and work through them collaboratively to reach consensus. It is an opportunity to step back once every five years to review what it is the unit should be doing and the extent to which it is achieving its aspirations.

14. Cyclical Review

Reviews are scheduled on a regular five year cycle, not in response to special needs or circumstances.

15. Complement Other Program Evaluations

Unit reviews are just one form of program and service evaluation, intended to draw upon and support other processes such as accreditation and curriculum review. Methods other than unit review should be used as circumstances warrant. These other forms of evaluation do not negate the need for unit review, but the unit review may be dramatically reduced in scope because information from these other sources can be used by the unit review committee.

16. Part of the Planning and Evaluation Cycle

Unit reviews should be linked to the College's strategic and tactical planning processes, should inform operational decision making, and should take into account findings from annual reports and institutional self-studies. Various follow-up reports are intended to facilitate these linkages.

17. Varying Levels of Reviews

Reviews are conducted with varying levels of comprehensiveness and thoroughness, specified in advance, according to the needs and circumstances of the unit. The more comprehensive reviews involve committees with additional members from outside the unit (chosen in consultation with the unit).